

PCT

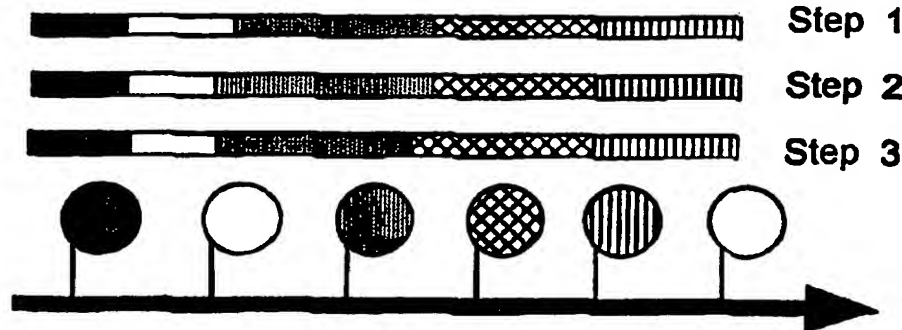
WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>H04L 12/43, 12/64</b>		<b>A1</b>	(11) International Publication Number: <b>WO 97/36402</b> (43) International Publication Date: <b>2 October 1997 (02.10.97)</b>
(21) International Application Number: <b>PCT/SE97/00520</b> (22) International Filing Date: <b>25 March 1997 (25.03.97)</b> (30) Priority Data: <b>9601130-9</b> <b>25 March 1996 (25.03.96)</b> <b>SE</b> (71) Applicant (for all designated States except US): <b>NET INSIGHT AB [SE/SE]; Ingenjörsvägen 3, S-117 43 Stockholm (SE).</b> (72) Inventors; and (75) Inventors/Applicants (for US only): <b>LINDGREN, Per [SE/SE]; Maria Prästgårdsgatan 12, S-118 52 Stockholm (SE). BOHM, Christer [SE/SE]; S:t Paulsgatan 28A, S-118 48 Stockholm (SE).</b> (74) Agent: <b>AWAPATENT AB; P.O. Box 45086, S-104 30 Stockholm (SE).</b>		(81) Designated States: <b>AL, AM, AT, AT (Utility model), AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, CZ (Utility model), DE, DE (Utility model), DK, DK (Utility model), EE, EE (Utility model), ES, FI, FI (Utility model), GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (Utility model), TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).</b>  <b>Published</b> <i>With international search report.</i>	

(54) Title: **REALLOCATION PROCEDURE**



(57) Abstract

The present invention relates to a method, a controller and a system for allocating access to time slots used for transferring data between nodes in a time multiplexed network, whose bandwidth is divided into cycles which in turn are divided into control slots for control signalling and data slots for transferring data, wherein slots allocated to a primary home node may be temporarily allocated to a temporary home node, said temporary home node instead of said primary home node then temporarily having the write access to said slots, and subsequently reallocated to said primary home node. According to the invention, the degree of temporary allocation of slots to a first node is evaluated and the number of slots having said first node as primary home node is increased, based upon said evaluation and in response thereto, in order to decrease the need for temporary allocation of slots within said network.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

REALLOCATION PROCEDURETechnical Field of Invention

The present invention relates to methods, a controller and a system for allocating access to time slots used for transferring data between nodes in a time multiplexed network, such as a Dynamic Synchronous Transfer Mode (DTM) network, whose bandwidth is divided into cycles which in turn are divided into control slots for control signalling and data slots for transferring data, wherein slots allocated to a primary home node may be temporarily allocated to a temporary home node, said temporary home node instead of said primary home node then temporarily having the write access to said slots, and subsequently reallocated to said primary home node.

15 Description of related art

As an example of a time multiplexed network, DTM is a broadband network architecture (see e.g., Christer Bohm, Per Lindgren, Lars Ramfelt, and Peter Sjödin, The DTM Gigabit Network, Journal of High Speed Networks, 3(2), pp 109-126, 1994 and Lars Gauffin, Lars Håkansson, and Björn Pehrson, Multi-gigabit networking based on DTM, Computer Networks and ISDN Systems, 24(2), pp 119-139, April 1992). DTM is a circuit switched network and intended to be used in public networks as well as in local area networks (LAN's). It uses channels as the communication abstraction. The channels differ from telephony circuits in various ways. First, establishment delay is short so that resources can be allocated/reallocated dynamically as fast as user requirements change. Second, they are simplex and so minimise overhead when the communication is unidirectional. Third, they offer multiple bit-rates to support large variations in user capacity requirements. Finally, they are multicast, allowing several destinations.

DTM channels share many beneficial properties with circuits. There is no transfer of control information after channel establishment, resulting in very high utilisation of network resources for large data transfers. Support of real-time traffic is natural; there is no need for policing, congestion control or flow-control within the network. The control information is separated from data, which makes multicast less complex. The transmission delay is negligible (i.e., less than 125 ms) and there is no potential for data loss because of buffer overflow as in ATM (Asynchronous Transfer Mode). Bit-error rates depend on the underlying link technologies, and switches are simple and fast due to strict reservation of resources at channel set-up. DTM can show good performance in areas where traditional circuit-switched networks fall short: dynamic bandwidth allocation, channel establishment delay, and as shared media networks.

The basic topology of a DTM network is preferably a bus with two unidirectional optical fibers connecting all nodes, but can also be realised by any other kind of structure, e.g., a ring structure. The DTM medium access protocol is a time-division multiplexing scheme. The bandwidth of the bus is divided into 125  $\mu$ s cycles, which in turn are divided into 64-bit time slots. The number of slots in a cycle thus depends on the network's bit-rate. The slots are divided into two groups, control slots and data slots. Control slots are generally static and used to carry messages for the network's internal operation. The data slots are used to transfer user data between the nodes.

Generally, in each network node there is a node controller, which controls the access to data slots and performs network management operations.

Control slots are used exclusively for control messages between these node controllers. Each node controller preferably has write access to at least one control

slot in each cycle, which it uses to multicast control messages to other node controllers. Here, multicast refers to sending information to one or more downstream nodes on a bus, as the transmission medium is unidirectional. Since write access to control slots is exclusive, the node controller always has access to its control slots regardless of other nodes and network load.

The majority of the slots in a cycle are data slots. Access to data slots changes over time, according to traffic demands. Write access to slots is controlled by slot access, sometimes referred to as slot tokens. A node controller may write data into a specific slot only if the node has write access to this specific slot. The slot access protocol, or token protocol, guarantees the slot access to be conflict free, which means that several nodes do not write data into the same slot.

One optimisation in slot access management is to introduce so called block access or block tokens, i.e. access to a number of slots arranged contiguously in a slot range. A block token, i.e. access to a group of slots, may be transferred in a single control message, but can only be used for particular combinations of slots. For instance, block access may be denoted by a slot number and an offset giving the number of contiguous slots in the group or block.

Also, write access to slots (single or block) may be exchanged or varied between different nodes during network operation, for example as a result of a node asking for more transferring capacity by requesting write access to more slots from other nodes. One of the major problems with this kind of write access allocation is fragmentation. The average number of contiguous available time slots allocated to a node of a DTM-like network is small due to the random movement and exchange of slot access and the varying capacity of users' requests. This gives relatively long access delays (in the millisecond region)

for high capacity channels, particularly at moderate to high load.

One way of dealing with this fragmentation problem is to define a home node for each slot at network start-up or during network operation. This is preferably done in such a way that slots having the same home node preferably define at least partly a continuous slot range. The nodes may then "borrow" available or free slots from each other, sending back free slots to their respective home node, for example when a significant time has passed. Also, two or more consecutive slots are preferably merged into a single block token when existing in the free pool of a node.

Hence, in a DTM-like system of this kind, a first node may have write access to a first set of slots and a second node may have write access to a second set of slots. Then, the first node is said to be the primary home node of the first set of slots and the second node is said to be the primary home node for the second set of slots. However, if the first node at some time require more transferring capacity to fulfil the demands of one or more users attached to the first node, it may temporarily lend or borrow slots from another node, for example the second node, presently having a surplus of slots. In such a case, one or more slots having the second node as primary home node may be temporarily allocated to the first node, the first node then being a so called temporary home node of said one or more borrowed slots and hence temporarily having the write access thereto. When the demand for transferring capacity in the first node decreases, or for example after a predefined time interval or, said one or more borrowed slots are returned to their primary or actual home node, i.e. the second node.

However, nodes in, e.g., an integrated services network will have different load depending on the equipment attached to it. Also, the load will vary in time. In case

of non-uniform traffic, a node that requires high capacity may have to regularly borrow slots from and return slots to their respective home node, which results in overhead for slot negotiation and longer set-up times for the channels.

Summary of the invention.

The object of the invention is to avoid the problems mentioned above with overhead for slot access negotiation in a time multiplexed network, such as a Dynamic Synchronous Transfer Mode (DTM) network, whose bandwidth is divided into cycles, which in turn are divided into time slots comprising control slots for control signalling and data slots for transferring data.

According to the invention, this and other objects are achieved by a method, a controller and a system having the features presented in the accompanying claims.

Hence, according to one aspect, the invention is preferably provided in a network wherein slots allocated to a primary home node may be temporarily allocated to a temporary home node, said temporary home node instead of said primary home node then temporarily having the write access to said slots, and subsequently reallocated to said primary home node.

According to the invention, the degree of temporary allocation of slots to a first node is evaluated and, based upon this evaluation, the number of slots having said first node as primary home node is increased in response thereto in order to decrease the need for temporary allocation of slots within said network.

Preferably, the transfer of slot ownership is initiated by a request from the node having a lack of slot capacity. Such a request, or any other type of reallocation initiation, may for example be provided when the node has borrowed slots a predefined number of times, when the ratio between the number of slots having the node as temporary home node and the number of slots

having the node as primary home node exceeds a predefined value, when the allocation or "loan" has to be directed over a substantial distance over the network. As another alternative, a predefined fraction or a predefined number of slots may be arranged to change owner, i.e. primary home node, each time one node borrows slots from another as a result of the loan itself.

Hence, according to the invention slots may be reallocated to new home nodes during network operation.

10 In one embodiment, a first node sends a request to a second node to get more slots assigned to it as primary home node and the second node transfers slots to the first node so that the slot capacity assigned to the first node as primary home node is increased and the slot

15 capacity assigned to the second node as primary home node is decreased. Preferably, the allocation is done in such a way that the tokens that will share the same primary home node will define at least partly a continuous slot range.

20 Preferably, slots having the same primary home node define a continuous slot range. Then, all nodes assigned as primary home nodes for slots located, in a cycle, in-between slots having a first node as primary home node and slots having a second node as primary home node may

25 transfer slot write access in the following way.

The second node is assigned as present node. Write access to a number of slots, amounting to the number requested by the first node, having the present node as primary home node and neighbouring a slot range having a

30 next node as primary home node are reallocated from the present node to the next node, said slot range of the next node either being the slot range of the first node or being a slot range of a node located in between the slot ranges of the first and the second nodes. The

35 distance between the slot range of the next node and the slot range of the first node is preferably smaller than the distance between the slot range of the present node



and the slot range of the first node. In case of the next node not being the first node, the next node is assigned as the present node and the step of reallocating is repeated. This procedure is referred to as a "push",  
5 since write access to the requested number of slots is pushed from the second node to a next node to another next node and so on, until it has been pushed to the first node.

According to the invention, the problems mentioned  
10 above are avoided by efficiently reallocating slot access to new primary home nodes, so minimising overhead for slot access negotiation. The criteria for when a node should request more slots assigned to it as primary home node, which node it should ask and how much capacity it  
15 should request, can be chosen in various ways, as mentioned above and below. However, resilience has to be implemented in order to avoid oscillation and thus overhead for control signalling.

Preferably, but not necessarily, each node has a  
20 node controller, which controls the write access to slots and in which a home node is defined for each slot and free slots may be sent back to their respective home nodes when a significant time has passed.

An advantage of the invention is that it is a simple  
25 and easily implemented mechanism, which strongly improves the performance of the network. The probability of that a node has to temporarily borrow slots from another node is decreased.

A further advantage of the invention is that low  
30 fragmentation is maintained.

Yet another advantage is that the reallocation of a slot to a new home node may be realised even though the slot is in use, e.g., by another node temporarily having access to the node, i.e. being the temporary home node of  
35 said slot.

Still another advantage of the invention is that it, especially in combination with a slot reuse method, may further improve the bandwidth utilisation.

5    Brief description of the drawings

Exemplifying embodiments of the invention will be described in greater detail below with reference to appended drawings, in which:

Fig 1 shows a dual-bus DTM network;

10    Fig 2 shows a DTM 125  $\mu$ s cycle that is divided into data slots and control slots;

Fig 3 shows different block slot access or block tokens in a slot-segment map;

15    Fig 4 shows how new home nodes are defined for slots in consecutive steps according to the invention; and

Fig 5-8 show results from simulations where the ownership reallocation and push mechanisms are included.

Detailed description of preferred embodiments

20    First, as an example of a time multiplexed network, the DTM MAC (Medium Access Control) protocol will be described with reference to Fig 1 and 2. The basic topology of a DTM-like network is a bus with two unidirectional optical fibers connecting all nodes Node 0, Node 1,  
25    ..., Node N-1, as shown in Fig 1. Several buses with different speeds may be connected to form an arbitrary multistage network. In the current prototype implementation, buses can be combined into a two-dimensional mesh. A node at the junction of two buses can synchronously  
30    switch data slots between the two buses. This allows for efficient switching with constant delay through the node. The primary communication abstraction in DTM is a multi-rate, multicast channel.

35    A DTM medium access protocol is a time-division multiplexing scheme. The bandwidth of the bus is divided into 125  $\mu$ s cycles, which in turn are divided into 64-bit time slots (or slots for short) as illustrated in Fig 2.

The number of slots in a cycle thus depends on the network's bit-rate; for instance, on a 6.4 Gbit/s network there are approximately 12500 slots per cycle.

As shown in Fig 2, the slots are further divided  
5 into two groups, control slots and data slots. Control slots are used to carry messages for the network's internal operation, such as messages for channel establishment and bandwidth reallocation. The data slots are used to transfer user data and are not necessarily read by  
10 intermediate network nodes, i.e. nodes arranged between the source node and the destination node or nodes.

Generally, but not necessarily, in each network node there is a node controller NC, which controls the access to data slots and performs network management operations,  
15 such as network start-up and error recovery. The main tasks of the node controller NC are to create and terminate channels on demand from users and to manage network resources in response to user requests and in the background.

20 Control slots are used exclusively for messages between node controllers NC. Each node controller NC preferably has write permission to at least one control slot in each cycle, which it uses to broadcast control messages downstream to other nodes. Since write access to  
25 control slots is exclusive, the node controller NC has always access to its control slots regardless of other nodes and network load. The number of control slots a node uses may vary during network operation.

The network is not restricted to a dual-bus, as  
30 exemplified in Fig 2, but can be realised by other kind of structures, e.g., a ring structure with an arbitrary number of nodes. The transmission media can besides to optical fibers be coaxial cables or any other high bandwidth transmission media. The bandwidth of the DTM dual  
35 bus in the preferred embodiment is divided into 125  $\mu$ s cycles, which in turn are divided into 64-bit time slots. The invention is of course not restricted to time

multiplexed networks with these values, but can be used in networks with cycles and slots of arbitrary sizes.

Principles for resource management (called slot access of token management) will be described below. The majority of the slots in a cycle are data slots. Access to data slots changes over time, according to traffic demands. Write access to slots is controlled by so called slot access. A node controller NC may write data into a slot only if the node has write access thereto. The token or write access protocol guarantees the slot access to be conflict free, which means that several nodes do not write data into the same slot. The write access to slots are controlled by the node controllers and it may be transferred from a node controller to another via control messages. In the following the write access protocol may also be refereed to as the token protocol.

Control messages for channel establishment and bandwidth reallocation contains set of slots or tokens as parameters. However, a control message is 64 bits and can therefore have only a small number of parameters. This means that if a user requests a large bandwidth transfer, it may be necessary to send several control messages to create the channel. This introduces extra access delay and consumes control signalling capacity.

Several mechanisms to decrease the amount of information that needs to be sent during channel creation and slot reallocation are considered. The first optimisation in token management is to introduce so called block tokens or slot range access. Access to a slot range may be transferred in a single control message and represents access to a group of slots, i.e. a group of slot tokens, but can only be used for particular combinations of slot access. For instance, block token may be denoted by a slot number and an offset giving the number of contiguous slots in the group or block.

In a basic form, the slot access protocol guarantees that a data slot can never be used by two nodes simulta-

neously on the bus, which sometimes may be too conservative. Nodes are connected by bus segments. Channels typically use a subset of the segments on the bus, and the rest are reserved but left unused and thus wasting shared resources. A better alternative is to let channels only reserve capacity on the segments between the sender node and the receiver node or nodes, as exemplified in Fig 3. A single slot may in this case be used multiple times on the bus. For example, channel D and E are using the same slots as channel A and C but on different segments. This is referred to as slot reuse. Slot reuse enables simultaneous transmissions in the same slot over disjointed segments of the bus. To allow slot reuse in a DTM-like network, the slot access format is extended to include parameters describing the segment(s) it is representing. The slot access management protocol is also modified to avoid conflicts in the slot number dimension as well as in the segment dimension.

A distributed slot access manager will be described in the following. Each node regularly broadcasts status information about how many free slots or slot ranges it has and how large they are. The other nodes store this information in their status tables. A node that wants more capacity consults its status table to decide from which node to request slots. The broadcast state information gives an approximate and dated view of the current state of slot information, so slot requests may be rejected because they were sent to nodes that no longer have free slots to give away.

The protocol on the initiating side may for example work as follows when a user request arrives to a node:

1. If the node has sufficiently many free slots to satisfy the request, it allocates the requested amount of slots to the user, and starts the channel by sending a channel establishment message to the destination node and then transmitting data using the reserved slots.

2. Otherwise the node marks its available slots as being reserved, and then preferably, but not necessarily, checks its status table: if the total amount of free slots in the network is not enough to fulfil the request, then the request is rejected (blocked). Otherwise the node requests slots from nodes with unused capacity.

3. If one of these nodes receiving a slot request does not have the requested amount of free slots, it may still give away all the slots it has. In any case, it sends a response back to the requesting node. A node preferably, but not necessarily, fulfils incoming requests in strict first-in-first-out order.

4. When a node receives a response to a slot request, it marks the slots it receives in the response (if any) as being reserved. When the node has received responses to all requests it has sent, it either starts the channels or rejects the user request, depending on whether or not it has acquired sufficient capacity. If the user request is rejected, the reserved slots are marked as free again.

At start-up all free slots may be evenly distributed among the network nodes, and each node may preferably then take at least one of its free tokens, move it (them) to the active state and declares it (them) to be a control slot. User requests may then be accepted and slot access can be moved between nodes on demand.

The pool of free slots may be distributed in other ways than evenly, e.g., proportionally (nodes upstream get more slots than nodes downstream) among all nodes.

In the general case, the average number of contiguous free slots in a node is small due to the random movement of slots and the varying capacity of users' requests. This fragmentation renders the block token optimisation practically useless, and the access delay is relatively long (milliseconds) for high capacity channels. To make block allocation efficient, it is necessary to reduce fragmentation of free slots or slot ranges, other-

wise fragmentation will be by far the main contributor to access delay for high bandwidth channels at moderate to high load. Low capacity channels will usually have a very short channel establishment delay independent of the  
5 current amount of fragmentation.

In the case of slot reuse, the fragmentation problem is even worse, as fragmentation may occur in both slot (time) and segment (space) dimensions (see Fig 3). In the distributed slot access manager system, most of the frag-  
10 mentation is a result of using many free pools (one for each node). Two free adjacent tokens, i.e. slots or slot ranges, can only merge if they are found in the same node.

An example of a distributed scheme that tries to  
15 avoid fragmentation if possible and increases the average block size of free slots belonging to a node works as follows:

1. A home node is defined for each slot or slot range at network start-up, and the slots are distributed  
20 in such a way so slots having the same primary home node will always define at least partly a continuous slot range. This results in a large average slot access area or token area in the token map shown in Fig 3.

2. When two consecutive slots or slot ranges having  
25 the same slot or segment range exist in the free pool, they are merged into a single token, i.e. a single slot range access (sometimes a recursive split and merge operation is needed). When performing a merge, a segment merge is prioritised before a slot number merge. (The  
30 reason for this is that slot access spanning over just a few segments are less useful to other nodes than slot access spanning over many segments.) Two segment-consecutive slots representing at least partly the same slot range and existing in the free pool of a node are split  
35 to obtain segment-consecutive slot access, representing the same slot range, which are merged into a single token.

3. When a node gets a slot access request from the local user or a remote user, slots is picked from the free slot pool, preferably, but not necessarily, using a best-fit algorithm in slot number and segment number dimension (see Fig 3). The "value of a token" is calculated as the area of a token in the token map and the token with the smallest area that fulfil the requested capacity is picked. A cost function can also be defined as a function of, e.g., number of slots, number of segments, location of slots and location of segments, which function should be minimised, however, still fulfilling the requested capacity.

4. When a node needs to request slots from other nodes, it does not ask for small chunks from several nodes if it is possible to ask for larger chunks from fewer nodes. The status tables provide this information. Transfer of slot access is therefore more efficient, and there are fewer establishment messages and less fragmentation.

5. Free slots are sent back to primary home nodes when they have been idle for a significant time or after a long transfer or after having been used a certain number of times or the like.

This scheme returns slots to their respective primary home nodes as a way to increase the probability that two consecutive slot or slot ranges can be merged in the free list, which decreases fragmentation. If the slots are returned to their respective primary home nodes too soon, i.e., if the home node "gravity" is too strong, the scheme will result in less sharing of resources and unnecessary control signalling. If it is too weak, fragmentation will still remain a problem. The "gravity" may be changed during the operation of the bus.

However, nodes in, e.g., an integrated services network will have different load depending on the equipment attached to it. Also, the load will vary in time. In case of non-uniform traffic, a node that requires high capaci-



ty may have to regularly return slots to their respective home nodes, which results in overhead for slot access negotiation and longer set-up times for the channels.

To avoid this, a new mechanism which defines new  
5 home nodes for slots during network operation, is introduced:

According to an exemplifying embodiment of the invention, a first node sends a request to a second node to get more slots assigned to it as primary home node and  
10 the second node transfers slots to the first node as new primary home node so that the slot capacity assigned to the first node as primary home node is increased and the slot capacity assigned to the second node as primary home node is decreased. Preferably this is done in such a way  
15 that the slots that will have the same primary home node will define at least partly a continuous slot range.

The invention avoids the problems with overhead for slot access negotiation by efficiently reallocate slots to new home nodes. Using the invention, low fragmentation  
20 is maintained. It is desirable to be able to define new primary home nodes for slots even though they are in use, i.e. even though they are currently allocated to some other node as temporary home node.

Preferably, slots sharing the same primary home node  
25 define a continuous slot range. Then, all nodes assigned as primary home nodes to slots located, in a cycle, in-between slots having the first node as home node and slots having the second node as home node may for example transfer slot access in the following way:

- 30 1. The second node is assigned as present node.
2. Access to a number of slots, amounting to the number requested by the first node, having the present node as home node and neighbouring a slot range having a next node as home node are reallocated from the present  
35 node to the next node, said slot range of the next node either being the slot range of the first node or being a slot range of a node located in between the slot ranges

of the first and the second nodes and the distance between the slot range of the next node and the slot range of the first node preferably being smaller than the distance between the slot range of the present node and the slot range of the first node.

3. In case of the next node not being the first node, the next node is assigned as the present node and the step 2 of reallocating is repeated.

This procedure is referred to as push.

To explain this procedure in detail, an example is given with reference to Fig 4. Suppose there are six nodes connected to a bus as shown in Fig 4. To each node is assigned a set of consecutive slots in such a way that the first slot range within the cycle is assigned to node 1 as home node, the second slot range within the cycle is assigned to node 2 as home and so on, as indicated in Step 1 of Fig 4.

Suppose now that node 4 has determined or evaluated that it often requests or borrows many slots and thus permanently needs more. It then requests, e.g., node 2 to transfer a number of slots in a slot range. The node to be requested can be chosen in several ways, as will be discussed below.

Node 2 now chooses a portion of its slots, which corresponds to a set of consecutive slots closest to the slot range assigned to node 3 as home node, and transfers it to node 3. Consequently, node 3 increases the size of the slot range assigned to it as home node, as shown in Step 2 of Fig 4.

Node 3 repeats this procedure and thus transfers a portion of its slot range, which is of the same size as the one received but corresponds to consecutive slots closest to the slot range assigned to node 4 as home node, to node 4. Now, the size of the slot range assigned to node 4 as home is increased, as indicated in Step 3 of Fig 4.

This method avoids fragmentation of the slot pools, since consecutive slots will be assigned to each node as home. Slot range access, or block tokens, which shall change home node, are in this manner "pushed" through the nodes between the two negotiating nodes (Step 1-3), thereof the name of the procedure.

With this configuration (node 1 is assigned as home to the first slot range, node 2 is assigned as home to the second slot range etc.) it is possible to implement a simple procedure for sending free slots, which may have been in use during the reallocation, to their respective new home nodes.

After the allocation, each home node has information about the slots assigned to it as home node. According to the simple procedure for sending free slots to their respective new home nodes, the node that uses the slot or slot range sends it back from where it got it when it does not need it any more. This because the node may not know the new home node after the reallocation. The node that receives the slot checks if the slot still is assigned to it as home node. If not, the node checks whether the slot or slot range corresponds to slot number(s) lower or higher than the slots corresponding to the slot range now assigned to it as home. If it corresponds to lower slot number(s), the slot or slot range access is sent to the node's closest neighbour with node number lower than its own, and if it corresponds to higher slot number(s), the slot or slot range access is sent to the node's closest neighbour with node number higher than its own. The next node that receives the slot or slot range access repeats the procedure until the slots are returned to its new home node. In this way, also returning slots are pushed through the nodes from the old home node to the new home node.

Using this simple procedure, the change of home node does not affect the use of the slots at that time. Since control messages are multicast in for example DTM, the

intermediate nodes can observe when nodes borrow slots. In this way the push procedure may be implemented in such a way that the intermediate nodes automatically perform the reallocations of new home nodes.

- 5        Which point in time a node should request more slots assigned to it as home may be decided every time the node requests or borrows slots from other nodes. To avoid oscillation, resilience has to be included, e.g., by requesting a home node change only every  $n$ 'th time a node
- 10       requests or borrows slots from other nodes or for example by only reallocating a certain part of the slots borrowed. Another technique that may be used is to correlate the node's borrowed slots to its owned. If, e.g., the quotient of borrowed slots, i.e. slots having the node as
- 15       temporary home node, and owned slots, i.e. slots having the node as primary home node, exceeds a predetermined value, a home node reallocation is initiated.

- From which node a home node movement shall be requested may be determined for example by using one or more
- 20       of the following methods.

- In the first method, the node requests the node which at the time has most free slots (capacity). This is easily found by referring to its status table, where information about the free capacity in each node is
- 25       stored.

- In the second method, the node maintains a record of nodes from which it has borrowed slots, e.g., most recently or during a predetermined period. It then requests a change of home node from the node it has,
- 30       according to the record, borrowed most slot capacity from.

- How many slots or how large slot range a node should request to be assigned to it as home may be chosen as a fraction of the slot capacity the node requests or
- 35       borrows from other nodes. It may be implemented as the value at the time for reallocation or as a mean value by referring to the record of loans. Alternatively, the

quotient of borrowed slots and owned slots may be used. It may also be set as a fixed amount of slots or may be controlled by management operations.

In some situations, the procedures may be utilised more efficiently on the DTM bus when combined with a slot reuse method as to be described in the following. When the first node sends a request to the second node to get more slots assigned to it as home node, it requests slot access corresponding only to predetermined segments of the network. The second node transfers slot access corresponding only to these segments to the first node, letting slot access, which correspond to the same slots but on different segments, be unchanged. This increases fragmentation, but will in some situations be advantageous.

Results from simulations performed, with the ownership reallocation and push mechanisms included, will be described in the following. In the simulations, reallocation of resources in the background to adapt to non-uniform traffic, is evaluated. The frame size is chosen to be 2400 slots, which corresponds to a communication link capacity of approximately 1.2 Gbit/s. For lower communication link capacity less processing and control signal capacity for each node are required. Thus, it is possible to simulate smaller transfers without trashing phenomena, i.e., very low throughput at high offered loads due to lack of signal capacity. Small transfers are used to more clearly observe the difference when including the push mechanism. The traffic in each node is exponentially distributed and a few nodes are designed as hot-spots, i.e., they have much more traffic in the simulations.

Initially, the home node assignment is evenly distributed among the nodes as can be seen in Fig 5. The change of home node is executed when a node has received a number of requests for more capacity from another node. Then it "pushes" write access to one slot through the

intermediate nodes to the requesting node. In the simulations this is done for every 10'th request, which introduces inertia in the system. This is important since the push mechanism and ownership reallocation mechanism shall  
5 cope with long term variations in load. The state at approximately 100 ms is shown in Fig 6. During the operation of the network, the home node is changed according to the load, i.e., the hot-spots will have more slots assigned to them as home.

10 The ownership allocation mechanism will reduce the amount of reallocations of slots for, e.g., client/server like traffic. In Fig 7 is shown the throughput and in Fig 8 the channel set-up time as a function of offered load with and without the ownership allocation and push  
15 mechanisms. Results are presented for transfers of 4 kB and 8 kB. It can be observed that the throughput is significantly higher with the ownership allocation and push mechanisms. Also, the channel set-up time is lower. This is the case since the hot-spots do not need to  
20 return slots to other nodes as frequently as with an evenly distributed home node assignment.

According to the invention, there is provided efficient reallocation of slot access to new home nodes, so minimising overhead for slot access negotiation. The  
25 criteria for when a node should request more slots assigned to it as home node, which node it should ask and how much capacity it should request, can be chosen in various ways, e.g., as described above. Requests of these types may also origin from the equipment attached to the  
30 node. It is possible to, e.g., manually initiate the re-allocation and decide how many slots that are to be re-allocated. Also, the allocation of slots may be controlled by a network controller controlling the allocation of  
35 slots to one or more nodes, such as a master node controlling the allocation of slots to a slave node. However, resilience has to be implemented in order to avoid oscillation and thus overhead for control signalling.

The procedures and mechanisms according to the invention are simple and easily implemented. The performance of the time multiplexed network is strongly improved and the fragmentation is maintained low. Also, the  
5 reallocation of a slot to a new home node may be realised even though the slot is in use.

As is understood by those skilled in the art, different variations and modifications of the invention may be made without departing from the scope of the  
10 invention, which is defined by the accompanying claims.

CLAIMS

1. A method for allocating access to time slots used for transferring data between nodes in a time multiplexed network, whose bandwidth is divided into cycles which in turn are divided into time slots, wherein a slot allocated to a primary home node may be temporarily allocated to a temporary home node and subsequently reallocated to the primary home node, said temporary home node instead of said primary home node then temporarily having the write access to said slot, said method comprising the steps of:

evaluating the degree of temporary allocation of slots, and  
altering, based upon said evaluation, the allocation of slots to primary home nodes in order to decrease the need for temporary allocation of slots within said network.

2. A method as claimed in claim 1, wherein said slots comprises control slots control slots for control signalling and data slots for transferring data.

3. A method as claimed in claim 1 or 2, wherein said evaluating and altering steps comprises:

evaluating the degree of temporary allocation of slots to a first node, and  
increasing, based upon said evaluation, the number of slots having said first node as primary home node in response thereto in order to decrease the need for temporary allocation of slots within said network.

4. A method as claimed in claim 1, 2 or 3, wherein said altering step comprises:

said first node sending a request to a second node, requesting the allocation of slots to the first node as primary home node; and



said second node allocating in response thereto, based upon an evaluation of it's own slot demand, slots having the second node as primary home node to said first node as new primary home node, thereby increasing the  
5 amount of slots having the first node as primary home node and decreasing the amount of slots having the second node as primary home node.

10 5. A method as claimed in claim in any one of the preceding claims, wherein said evaluating step is performed by said second node.

15 6. A method as claimed in claim 1 or 2, wherein said evaluating and slot altering step is performed by a network control unit.

20 7. A method as claimed in any one of the preceding claims, wherein the step of altering or increasing of the number of slots having the first node as primary home node comprises primarily allocating slots thereto that are contiguous with other slots already having the first node as primary home node.

25 8. A method as claimed in any one of the preceding claims, said method being implemented in a Dynamic Synchronous Transfer Mode (DTM) network.

30 9. A method as claimed in any one of the preceding claims, wherein the slots to be allocated from a second node to the first node as primary home node are contiguous and thus define a slot range.

35 10. A method as claimed in any one of the preceding claims, wherein the allocating of a number of slots from a second node to the first node comprises:  
designating the second node as present node;

allocating said number of slots from the present node to a next node being a home node for a neighbouring slot or slot range; and

designating said next node as the present node and  
5 repeating the step of allocating said number of slots from the present node to a next node until said number of slots have been allocated to the first node.

11. A method as claimed in claim 10, wherein the  
10 number of slots between the slot or slot range of the next node and the slot or slot range of the first node is smaller than the number of slots between the slot or slot range of the present node and the slot or slot range of the first node.

15

12. A method as claimed in any one of the preceding claims, comprising the step of, at least during network start-up, assigning a contiguous slot range to a first node as primary home node and assigning another,  
20 preferably following, contiguous slot range to a next node as primary home node and so on, wherein the number of slot ranges preferably equals the number of nodes.

13. A method as claimed in claim 12, further comprising:  
25

informing each node about upper and lower bounds of the slot range having the respective node as primary home node;

30 comparing in a node the slot number of a slot with the upper and lower bounds of the slot range having the node as primary home node;

allocating, if said slot number is lower than the lower bound of the slot range having said node as primary home node, the slot to a next node having a neighbouring  
35 lower slot range;

allocating, if said slot number is higher than the upper bound of the slot range having said node as primary

home node, the slot to a next node having a neighbouring higher slot range.

14. A method as claimed in any one of the preceding  
5 claims, wherein said altering step is performed, as a result of said evaluating step, when a node has temporarily borrowed slots from other nodes a significant number of times.

10 15. A method as claimed in any one of the preceding claims, wherein said altering step is performed, as a result of said evaluating step, when the quotient of temporarily borrowed slot capacity of a node as temporary  
15 home node and the slot capacity assigned to said node as primary home node exceeds a predetermined value.

16. A method as claimed in any one of the preceding claims, wherein a first node sends a request to a second node, requesting the allocation of slots to the first  
20 node as primary home node, comprising selecting the second node as the node that has most free slot capacity.

17. A method as claimed in any one of the preceding claims, wherein a first node sends a request to a second  
25 node, requesting the allocation of slots to the first node as primary home node, wherein said second node is derived from a status table, which is maintained, preferably in each node, with information about the free capacity in respective node.

30 18. A method as claimed in any one of the preceding claims, wherein a first node sends a request to a second node, requesting the allocation of slots to the first node as primary home node, wherein the second node is  
35 selected as the node from which, during a predetermined period, the first node has requested or temporarily borrowed most slot capacity.

19. A method as claimed in any one of the preceding claims, wherein a first node sends a request to a second node, requesting the allocation of slots to the first node as primary home node, wherein said second node is derived from a record that is maintained, preferably in each node, with information about its requested or temporarily borrowed slot capacity during a predetermined period.

10

20. A method as claimed in any one of the preceding claims, further comprising, when a node requests a temporary allocation of slots to it as temporary home node, allocating a fraction, or at least one, of said slots to said node as primary home node and remaining slots to said node as temporary home node.

21. A method as claimed in any one of the preceding claims, wherein slots being allocated a first node as primary home node only refers to one or more predetermined segments of the network, thereby allowing the same slots to be allocated to other nodes on other segments.

22. A method for allocating access to time slots used for transferring data between nodes in a time multiplexed network, whose bandwidth is divided into cycles which in turn are divided into control slots for control signalling and data slots for transferring data, wherein a first node has write access to at least a first slot or slot range, comprising the steps of  
determining an increased demand for transferring capacity in the first node;

allocating an extra slot or slot range to the first node in such a way that said extra slot or slot range are contiguous with said first slot or slot range as a result of said increased demand.

23. A method as claimed in claim 21, wherein the allocating of a number of slots in a slot or slot range from a second node to the first node comprises:
- designating the second node as present node;
  - 5 allocating said number of slots from the present node to a next node having access to a neighbouring slot or slot range; and
  - designating said next node as the present node and repeating the step of allocating said number of slots
  - 10 from the present node to a next node until said number of slots have been allocated to the first node.
24. A method as claimed in claim 22, wherein the number of slots between the slot or slot range of the
- 15 next node and the slot or slot range of the first node is smaller than the number of slots between the slot or slot range of the present node and the slot or slot range of the first node.
- 20 25. A controller in a time multiplexed network, whose bandwidth is divided into cycles which in turn are divided into time slots, wherein slots allocated to a primary home node may be temporarily allocated to a temporary home node, said temporary home node instead of
- 25 said primary home node then temporarily having the write access to said slots, and subsequently reallocated to said primary home node, said controller comprising:
- means for evaluating the degree of temporary allocation of slots to a first node, and
  - 30 means for changing, based upon said evaluation, the number of slots having said first node as primary home node in response thereto in order to decrease the need for temporary allocation of slots within said network.
- 35 26. A controller as claimed in claim 25, wherein said slot comprises control slots for control signalling and data slots for transferring data.

27. A controller as claimed in claim 25 or 26,  
further comprising means for sending a request to a  
second node, requesting the allocation of slots to the  
5 first node as primary home node.

28. A controller as claimed in claim 25, 26 or 27,  
wherein said controller is arranged to primarily allocate  
slots to the first node that are contiguous with other  
10 slots already having the first node as primary home node.

29. A controller as claimed in claim 25, 26, 27 or  
28, wherein said controller is provided as a node  
controller in said first node.

15

30. A controller as claimed in anyh one of claims  
25-29, wherein said network is a Dynamic Synchronous  
Transfer Mode (DTM) network.

20 31. A system implemented in a time multiplexed  
network, whose bandwidth is divided into cycles which in  
turn are divided into time slots, wherein slots allocated  
to a primary home node may be temporarily allocated to a  
temporary home node, said temporary home node instead of  
25 said primary home node then temporarily having the write  
access to said slots, and subsequently reallocated to  
said primary home node, said system comprising:

means for evaluating the degree of temporary  
allocation of slots, and

30 means for changing, based upon said evaluation, the  
allocation of slots to different nodes as primary home  
nodes in response thereto in order to decrease the need  
for temporary allocation of slots within said network.

35 32. A system as claimed in claim 31, wherein each  
cycle are divided control slots for control signalling  
and data slots for transferring data.

33. A system as claimed in claim 31 or 32, further comprising:

means for sending a request to a second node,  
5 requesting the allocation of slots to the first node as primary home node; and

means for allocating, based upon an evaluation of the second nodes own slot demand, slots having the second node as primary home node to said first node as new  
10 primary home node, thereby increasing the amount of slots having the first node as primary home node and decreasing the amount of slots having the second node as primary home node.

15 34. A system as claimed in claim 31, 32 or 33, wherein said system is arranged to primarily allocate slots to the first node that are contiguous with other slots already having the first node as primary home node.

20 35. System as claimed in claim 31, 32 33 or 34, wherein a number of slots in a slot or slot range are allocated from a second node to the first node as home node, further comprising:

means for designating the second node as present  
25 node;

means for allocating said number of slots from the present node to a next node having access to a neighbouring slot or slot range; and

means for designating said next node as the present  
30 node and repeating the step of allocating said number of slots from the present node to a next node until said number of slots have been allocated to the first node.

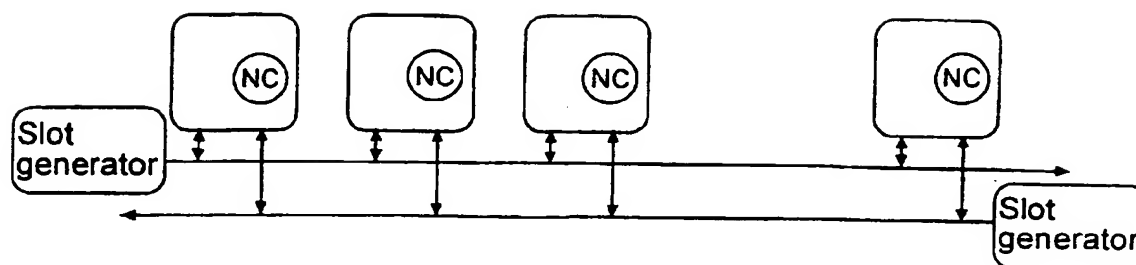


Fig. 1

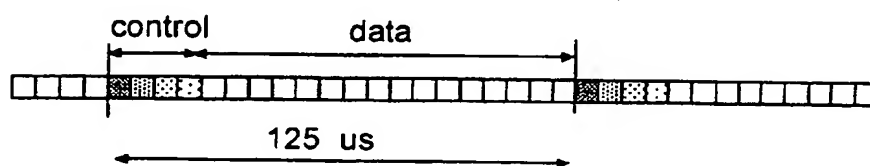


Fig. 2

**SUBSTITUTE SHEET**



2/4

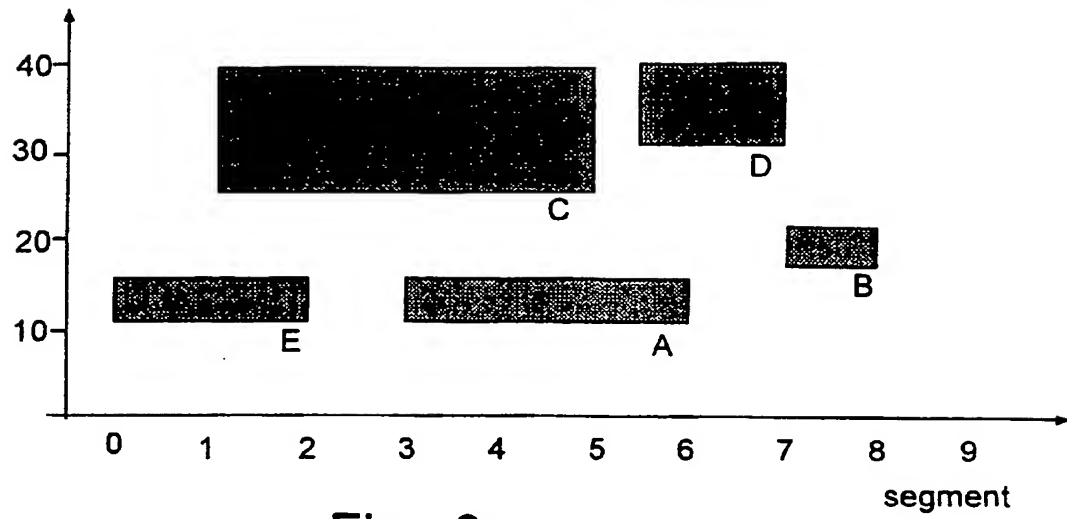


Fig. 3

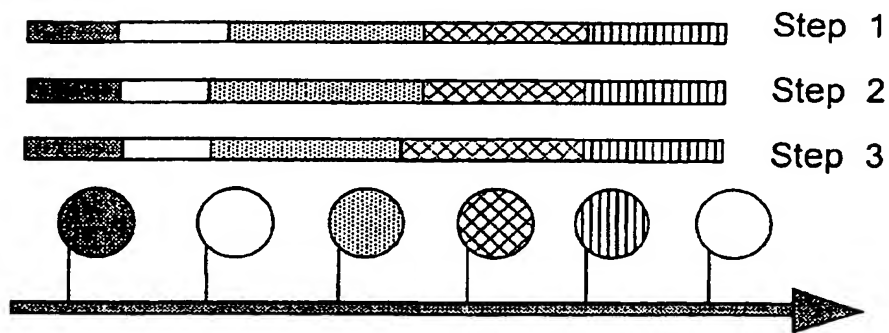


Fig. 4

**SUBSTITUTE SHEET**

Number of blocks assigned as home

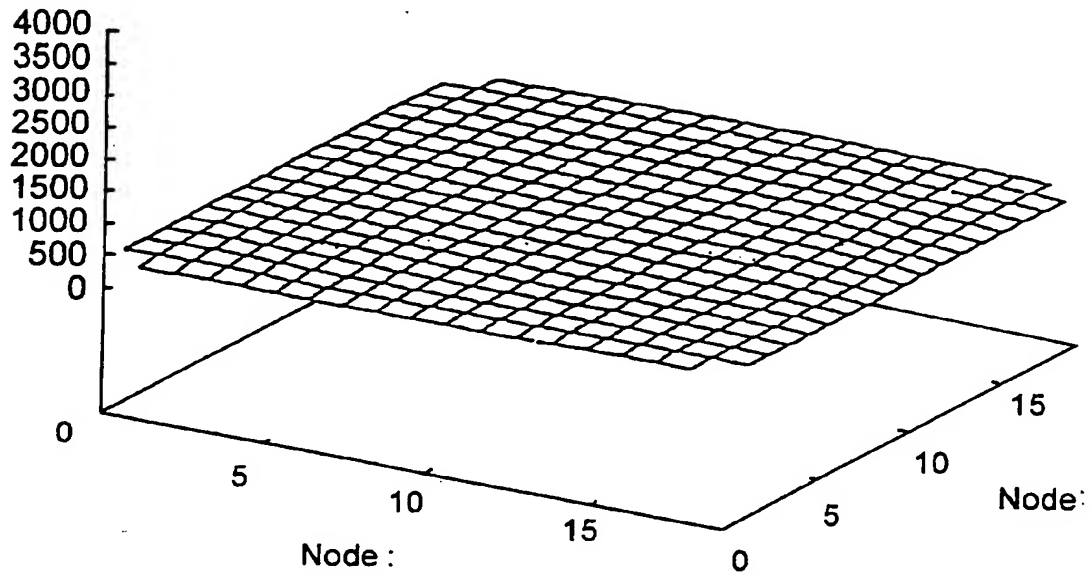


Fig. 5

Number of blocks assigned  
as home

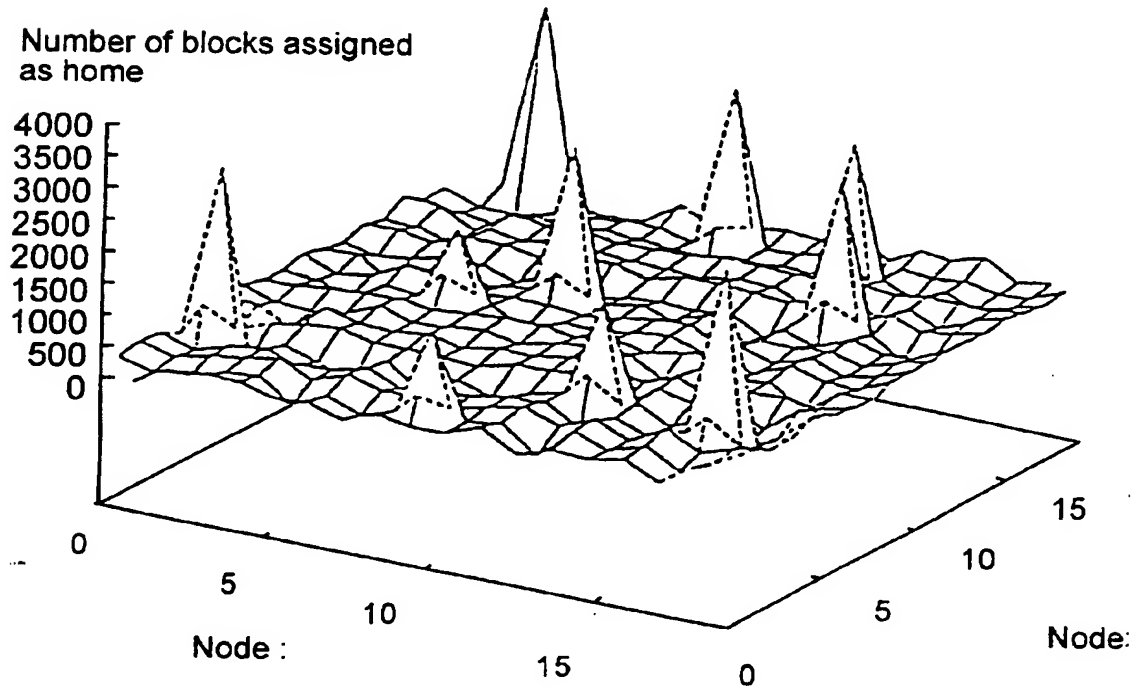


Fig. 6

**SUBSTITUTE SHEET**

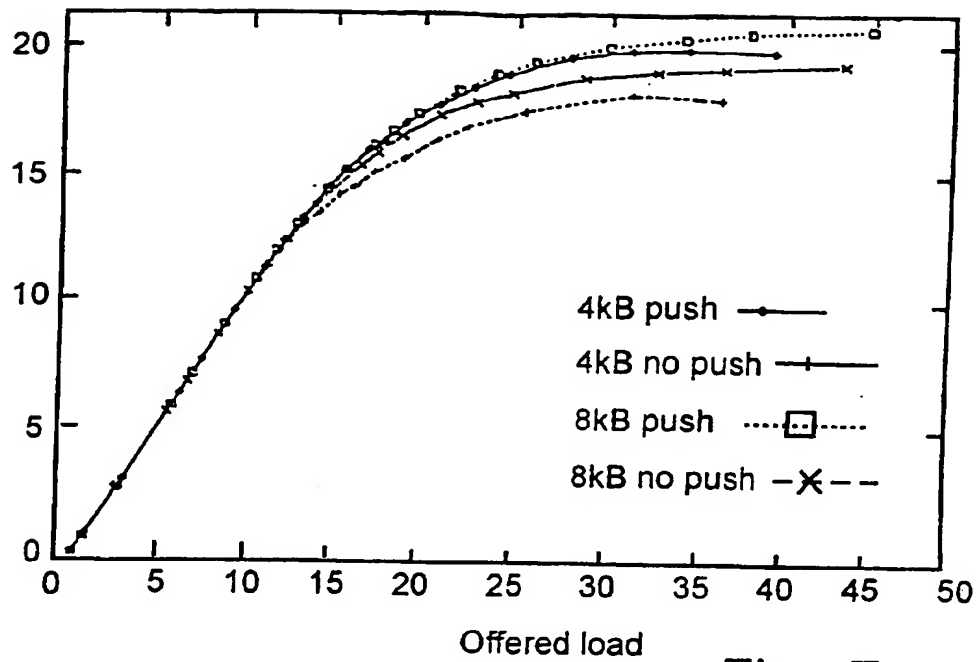


Fig. 7

Channel set-up time (microseconds)

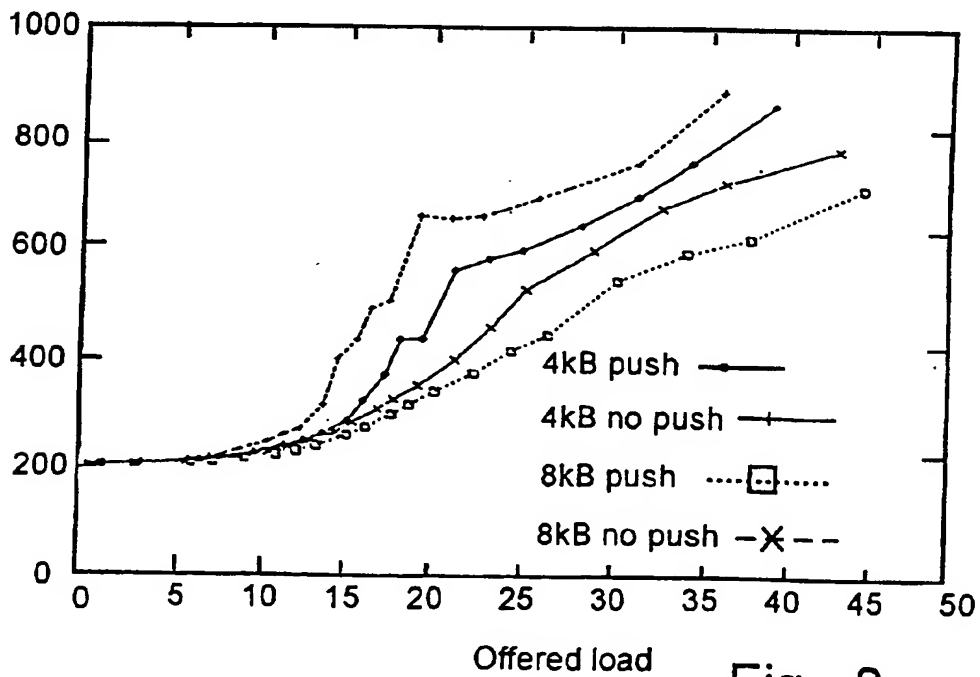


Fig. 8

**SUBSTITUTE SHEET**

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/SE 97/00520

<b>A. CLASSIFICATION OF SUBJECT MATTER</b>		
IPC6: H04L 12/43, H04L 12/64 According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols)		
IPC6: H04L, H04J		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
SE,DK,FI,NO classes as above		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
WPI, JAPIO		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	IEEE Journal on Selected Areas in Communications, Volume 14, No 2, February 1996, Bohm C. et al, "Fast Circuit Switching for the Next Generation of High Performance Networks" page 298 - page 305 --	1,2,4-8,10, 12,20,22, 25-35
X	Patent Abstracts of Japan, Vol 17, No 150, E-1339, abstract of JP,A,4-316240 (FUJI XEROX CO LTD.), 6 November 1992 (06.11.92) --	1
A	EP 0477659 A1 (ALCATEL FACE S.P.A.), 1 April 1992 (01.04.92), column 6, line 11 - line 34 --	1-35
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
8 July 1997		12.07.97
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86		Authorized officer  Friedrich Kühn Telephone No. +46 8 782 25 00

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/SE 97/00520

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 4949395 A (NILS R.C. RYDBECK), 14 August 1990 (14.08.90), column 8, line 1 - column 9, line 27 ---	1-35
P,A	EP 0726661 A2 (NEC CORPORATION), 14 August 1996 (14.08.96), abstract -- -----	1-35

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

01/07/97

International application No.  
PCT/SE 97/00520

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0477659 A1	01/04/92	SE 0477659 T3 AT 131327 T AU 635157 B AU 8458491 A DE 69115202 D,T ES 2083497 T IT 220082 Z IT 1243085 B JP 4264845 A US 5208808 A	15/12/95 11/03/93 02/04/92 18/07/96 16/04/96 09/06/93 23/05/94 21/09/92 04/05/93
US 4949395 A	14/08/90	AU 634792 B AU 5872390 A CA 2020625 A,C DE 69030534 D EP 0407367 A,B JP 3046424 A	04/03/93 10/01/91 08/01/91 00/00/00 09/01/91 27/02/91
EP 0726661 A2	14/08/96	JP 8204673 A	09/08/96